



UNITED STATES DEPARTMENT OF COMMERCE
Bureau of the Census
Washington, DC 20233-0001

September 22, 1999

DSSD CENSUS 2000 PROCEDURES AND OPERATIONS MEMORANDUM SERIES R-18

MEMORANDUM FOR Howard Hogan
 Chief, Decennial Statistical Studies Division

From: Donna Kostanich *DK*
 Assistant Division Chief, Sampling and Estimation
 Decennial Statistical Studies Division

Prepared by: James Farber *DF for JF*
 Sample Design Team

Subject: Accuracy and Coverage Evaluation Survey: Sample Reduction
 Overview

I. INTRODUCTION

The purpose of this document is to present an overview of the Accuracy and Coverage Evaluation (A.C.E.) sample reduction. The goals of the A.C.E. reduction are three-fold. First, the number of block clusters must be reduced from the listing sample size. The listing sample was selected under the previous 750,000 housing unit design, while the A.C.E. is a 300,000 housing unit design. Second, in reducing the block cluster sample size, the sample sizes for population subgroups that have historically been undercounted in the census should be increased relative to other subgroups to achieve reliable A.C.E. population estimates. The details of differential sampling by demographic group are still in research. The last goal of the A.C.E. reduction is to reduce the variance contribution from block clusters that may potentially be outlier clusters and thus exert undue influence on the population estimates. These clusters are identified by comparing their census and A.C.E. listing housing unit counts. The details of differential sampling for these clusters is also currently in research.

The A.C.E. reduction is one of the processes involved in determining which housing units will be interviewed in the A.C.E. survey. The A.C.E. reduction is a subsample of the medium (3 to 79 housing units) and large (80+ housing units) block clusters previously selected for the A.C.E. listing sample. Following the A.C.E. reduction, the small block cluster subsampling and large block cluster subsampling operations occur. The number of small block clusters (0 to 2 housing units) is reduced during small block cluster subsampling, and the number of housing units in large block clusters is reduced in large block cluster subsampling. The A.C.E. interview sample consists of those housing units in block clusters or segments of block clusters selected for interview in the A.C.E. reduction and subsequent subsampling processes.

II. A.C.E. REDUCTION DESIGN

The following are features of the planned design for the A.C.E. reduction:

- The A.C.E. reduction, along with small block cluster subsampling and large block cluster subsampling, is designed to achieve the national interview housing unit sample size of approximately 300,000 housing units.
- The A.C.E. sample is designed separately by state with the national sample allocated to states proportional to population with a minimum sample size of 1800 housing units.¹
- Block clusters in Puerto Rico will not be subsampled in the A.C.E. reduction. All Puerto Rico clusters will be retained in the A.C.E. interview sample.
- Block clusters on American Indian Reservations² will not be subsampled in the A.C.E. reduction. All American Indian Reservation clusters will be retained in the A.C.E. interview sample.

Clusters on Tribal Jurisdiction Statistical Areas, Tribal Designated Statistical Areas, and Alaska Native Village Statistical Areas probably will be subsampled in the A.C.E. reduction. This issue is currently unresolved.

- Small block clusters will not be subsampled in the A.C.E. reduction. A separate subsampling operation will reduce the number of small block clusters in the A.C.E. interview sample.

¹ Mule (June, 1999), "Accuracy and Coverage Evaluation Survey: State Interviewing Sample Size Estimates," DSSD Census 2000 Procedures and Operations Memorandum Series R.

² The American Indian Reservations includes the associated Trustlands.

- Only medium and large block clusters that are not on an American Indian Reservation and not in Puerto Rico are subsampled in the A.C.E. reduction.
- The calculation of reduction sampling rates is based on the most recent measure of size, the preliminary A.C.E. independent listing housing unit count. These housing unit counts are preliminary because the number is simply a clerical tally of the number of housing units listed in the independent listing book.
- Block clusters that were in the medium stratum at the time of listing sample selection but have 80 or more housing units based on the preliminary listing housing unit count will likely be retained at higher rates to control their weights. In the listing sample, medium clusters were sampled at lower rates than large clusters since large clusters eventually undergo large block subsampling, an operation that increases weights.
- Excluding medium clusters that have 80 or more housing units on the independent list, medium and large block clusters will be subsampled in the A.C.E. reduction at the same relative rates used in listing sample selection. That is, the differential allocation of medium and large clusters in the listing sample will be retained in the reduced sample.

III. RESEARCH ITEMS

The following issues for the A.C.E. reduction are still in research:

- Differential subsampling rates may be used for certain demographic groups, such as minority/non-minority, in states where the population is estimated to be sufficiently heterogeneous and where the listing sample size is sufficiently large. Differential sampling as opposed to proportional sampling could provide more reliable A.C.E. estimates for demographic groups that have historically been undercounted. Research is ongoing to determine whether differential sampling by demographic group is expected to provide variance reduction, and if so, what the differential subsampling rates should be to maximize variance reduction while also controlling weight variation. It is expected that no more than two demographic strata would be formed in a single state to control weight variation.
- Differential subsampling rates may also be used for clusters where the current census housing unit count differs significantly from the A.C.E. independent listing housing unit count. Clusters with significant differences are called “Inconsistent” while other clusters are “Consistent.” It is expected that only two strata will be formed, although it is possible that the Inconsistent stratum might be split into two

strata depending on the results of research. The definition of a significant difference is unresolved at this point. Possibilities include measures based on absolute or percent differences in the two housing unit counts. Inconsistent clusters are more likely to experience coverage problems and thus should be retained in the A.C.E. interview sample at a higher rate than Consistent clusters. The extent to which Inconsistent block clusters might be differentially sampled is currently unknown. All List Enumerate clusters will likely be considered inconsistent since the census housing unit count in these clusters is unknown at the time of the A.C.E. reduction.

- In many states, it is possible that differential sampling will be used based on both housing unit count differences and demographic groups, and thus these two types of strata need to be integrated in the A.C.E. reduction. To control weight variation, the current plan is to combine these two types of strata into three A.C.E. reduction strata:
 - Minority Block Clusters
 - Non-Minority Inconsistent Clusters
 - All Remaining Clusters

Alternative combinations are also under consideration. To further reduce weight variation, the first two A.C.E. reduction strata may possibly be given the same sampling rate. Other alternatives to control weight variation are also being researched.

- The plans are to use differential sampling conservatively. The research may show large variance gains by allowing extensive weight variation; these gains may not be achieved for 2000. From the demographic groups perspective, population shifts have probably occurred since 1990 which could cause wide weight variation within demographic groups that we would like to avoid. For the consistency of housing unit counts, the reduction of variance may not be as significant if the targetted extended search program is successful. Further, the relationship of the two housing unit counts is a proxy variable. Even when the two counts are relatively comparable for a cluster, it is possible for there to be coverage problems in the cluster.

cc: Census 2000 Procedures and Operations Memorandum Series Distribution List
Statistical Design Team Leaders
Sampling and Estimation Staff